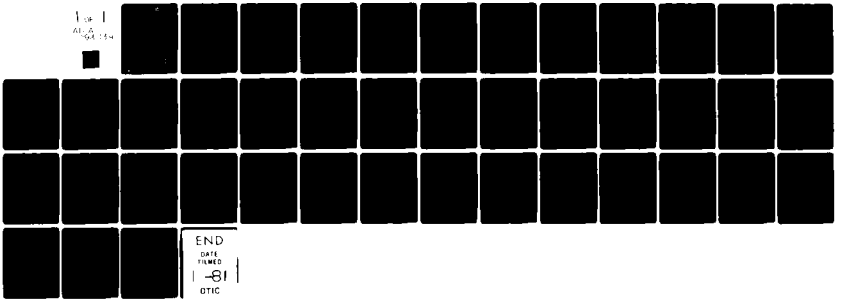
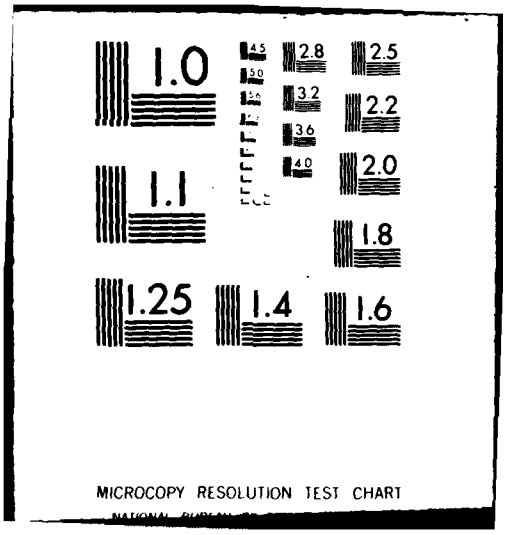
AD-A093 039 PRINCETON UNIV NJ DEPT OF STATISTICS F/G 12/1
NEW MODELS FOR RELIABILITY GROWTH.(U)
OCT 80 H BRAUN, N SCHENKER F49620-79-C-0148
UNCLASSIFIED TR-174 AFOSR-TR-80-1217 NL



END
DATE FILMED
-81
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART

AFOSR-TR- 80 - 1217

New Models for Reliability Growth

FINAL

by

Henry Braun*
Department of Statistics
Princeton University

Nathaniel Schenker**
Department of Statistics
Princeton University

October 1980

Technical Report No. 174

Research Sponsored by Air Force Office of Scientific Research Contract F49620-79-C-0148, Department of Statistics, Princeton University, Princeton, New Jersey.

*Currently at Educational Testing Service, Princeton, NJ.

**Currently at Department of Statistics, University of Chicago.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

(18) 1. REPORT NUMBER: AFOSR-TR-80-1217	2. GOVT ACCESSION NO.: AD-A093 039
3. RECIPIENT'S CATALOG NUMBER: (14) TR-1174	
4. TITLE (and Subtitle): (6) NEW MODELS FOR RELIABILITY GROWTH.	5. TYPE OF REPORT & PERIOD COVERED: (9) Final rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): (10) Henry Braun Nathaniel Schenker	8. CONTRACT OR GRANT NUMBER(s): (15) F49620-79-C-0148 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Princeton University, Department of Statistics, Princeton, New Jersey 08541	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: (16) 61102F 2304/A5 (17) A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, Bolling AFB, Washington, DC 20332	12. REPORT DATE: October 1980 (11)
	13. NUMBER OF PAGES: 40 (12) 43
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Reliability growth modeling is an important component in development testing of complex systems. It provides estimates of the current status of the process as well as forecasts of reliability at future time points. Several models have been investigated in the literature. In this report, we focus on comparing the performance of a model introduced by L. Crow and a family of models which are generalizations of those introduced in Braun and Paine. The novel feature of this latter family is that is models the logarithm of the intensity function of the failure process as a multiple linear regression on some function of the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406873

number of different failure modes uncovered and the number of recurrent failures. These models possess great flexibility in accomodating many different shapes of reliability growth curves including nonmonotonic ones. Three members of this family are compared with the Crow model over a range of simulated reliability growth patterns. One member, in particular, seems to provide short-and long term forecasts superior to those of Crow's. These results suggest that these new models be in the modeling of field data. Three such data sets are considered here as well.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

1. Introduction

Reliability growth modeling is an important component in development testing of complex systems. It provides estimates of the current status of the process as well as forecasts of reliability at future time points. Several models have been investigated in the literature (see Schafer, et al. [1975]). In this report, we focus on comparing the performance of a model introduced by L. Crow [1975] and a family of models which are generalizations of those introduced in Braun and Paine [1977].

The novel feature of this latter family is that it models the logarithm of the intensity function of the failure process as a multiple linear regression on some functions of the number of different failure modes uncovered and the number of recurrent failures. These models possess great flexibility in accommodating many different shapes of reliability growth curves including nonmonotonic ones.

Three members of this family are compared with the Crow model over a range of simulated reliability growth patterns. One member, in particular, seems to provide short-and long-term forecasts superior to those of Crow's. These results suggest that these new models be considered in the modeling of field data. Three such data sets are considered here as well. Throughout this report, the following notation will be employed.

Suppose a failure process consists of n failures. Let

t = operating time,

$N(t)$ = number of failures in $(0,t]$,

t_i = time of the i^{th} failure, $i=1,\ldots,n$,

$X_1 = t_1$ = first interfailure time,

$X_i = t_i - t_{i-1}$ = ith interfailure time, $i=2,\ldots,n$, and

$CMTBF(t) = t/N(t)$ = cumulative mean time between failure at time t.

2. Reliability Growth Models

A. Crow's Model

In 1964, J. T. Duane proposed a deterministic model for reliability growth in which he hypothesized that there is a linear relationship between ln(CMTBF(t)) and ln(t). A stochastic analog to this model was suggested by Crow [1975]. In this stochastic model, the failure process is a nonhomogeneous Poisson process with intensity function given by

$$h_c(t) = kbt^{b-1}, \quad k,b,t>0.$$

The maximum likelihood estimates (MLE's), $\hat{k}$ and $\hat{b}$, of k and b, are easy to compute (see Appendix). Using the MLE's, we have the following estimate for the intensity function:

$$\hat{h}_c(t) = \hat{k}\hat{b}t^{\hat{b}-1}.$$

If $h_c(t)$ were constant between failures, then each X_i would have an exponential distribution with parameter $h_c(t_{i-1})$. It would, therefore, be true that $E(X_i)=1/h_c(t_{i-1})$. Although Crow's intensity function is not constant between failures, the change in the intensity from one failure to the next is small. Therefore, $1/h_c(t_{i-1})$ is very close in value to $E(X_i)$. Throughout this paper, we will use

$$\hat{E}(X_i) = 1/\hat{h}_c(t_{i-1})$$

as an estimate of $E(X_i)$ for Crow's model.

The intensity $h_c(t)$ in Crow's model is a monotone function. It decreases when $b<1$, increases when $b>1$, and is constant when $b=1$. These three possibilities correspond, respectively, to reliability growth, reliability loss, and constant reliability. Suppose a process exhibits neither strict growth nor strict loss of reliability. For example, a failure process may at first display a loss of reliability followed by reliability growth. In this case, the plot of $\ln(\text{CMTBF}(t))$ versus $\ln(t)$ will look U-shaped. This phenomenon corresponds to an intensity function which at first increases and then decreases. For such data, a model with a monotone intensity function will not be entirely suitable.

Another possible limitation of Crow's model is that it does not distinguish between failure types. In many applications, the data consist not only of the time of failure but also the (probable) cause. Thus, it is known whether a new failure mode has been revealed or whether an old one has recurred. Presumably, for a fixed number of observed failures, different combinations of new and recurrent failure modes are associated with different patterns of future reliability growth.

B. Failure Mode Models

In order to circumvent the above-mentioned constraints in the Crow model, a more flexible family of models has been suggested by the first author. The principal aim of this investigation has been to determine the properties of this new family and compare its performance to that of Crow's.

Let, then,

$N(t)$ = number of failures in $(o,t]$;

$N_1(t)$ = number of different failure modes discovered in $(o,t]$;

$N_2(t) = N(t) - N_1(t)$

Thus, $N_2(t)$ gives the number of failures in $(o,t]$ which are recurrences of previously observed modes. It will be assumed that the failure process is a Poisson process with intensity function $h_f(t)$ defined by

$\ln(h_f(t)) = d_o + d_1 \ln(N_1(t) + 1) + d_2\, g(N_2(t) + 1)$, where d_o, d_1, d_2 are parameters and $g(\cdot)$ is a function to be chosen. Here, we shall consider three candidates for $g(\cdot)$: the identity, the square root, and the natural logarithm. The MLE's for d_o, d_1, and d_2 can be obtained by iterative methods (see Appendix).

Note that $h_f(\cdot)$ is a step-function in time with jumps occurring at each point of failure. Aside from some technical points, the failure process is an inhomogeneous Poisson process in which each interfailure time X_i has exactly an exponential distribution with parameter $h_f(t_{i-1})$. Since $E(X_i) = 1/h_f(t_{i-1})$, the estimate $\hat{E}(X_i) = 1/\hat{h}_f(t_{i-1})$ will be used throughout.

The above family does have the flexibility to accommodate nonmonotone intensity functions. For example, imagine a system for which in the early stages of reliability growth testing $N_1(t)$ increases more rapidly than $N_2(t)$. Once most of the failure modes have been discovered, though not entirely corrected, $N_1(t)$ will increase only infrequently in comparison to $N_2(t)$. If $d_1 > o$ and $d_2 < o$, then $h_f(t)$ will increase at first and subsequently decrease. On the other hand, if d_1 and d_2 are both negative, then $h_f(t)$ will be monotone decreasing.

The family of failure mode models is very much empirical in nature. It is based on the assumption that the patterns in $N_1(t)$ and $N_2(t)$ jointly provide more information on the structure of reliability growth than N(t) alone. Under this assumption it seems more natural to calibrate time in terms of $N_1(t)$ and $N_2(t)$. This permits the model to deal with systems in which the rate of reliability growth is nonmonotone. Of course, the validity and utility of this point of view must be borne out in practice.

C. Urn Models

We also investigated the possibility of using urn models to represent the failure process. Imagine an urn containing (C + N) balls: C clear balls and N colored ones. Balls are drawn from the urn at the rate of one every δ hours. If a clear ball is drawn, no failure is recorded and the ball is returned to the urn. If a colored ball is drawn, a failure is recorded and a clear ball is returned to the urn in place of the ball drawn.

Under this sampling scheme,

$$E_{N_1,C}(X_n) = \delta(N+C)/(N-n)$$

and

$$E(Y_{n+1}) = \delta(C+N)\ [(n/N) + 1/2\ (n/N)^2].$$

Maximum likelihood estimates for the parameters can be obtained by grid search methods and the derivations are presented in the Appendix. Unfortunately, there are occasional difficulties with convergence and the estimates seem to be exceedingly variable.

3. Data

A. Field Data

We have available for analysis three data sets obtained from actual development testing. The first set consists of 52 failure times of a complex electronic system built by General Electric. Aside from 14 so-called non-pattern failures, all other failures were associated with one of 13 identifiable failure modes. In one version of this data set, to be denoted by GE, each non-pattern failure is treated as representing a newly discovered failure mode - making 27 in all. In another version, denoted by GE, the non-pattern failures are treated as recurrences of a single failure mode - making 14 failure modes in all.

The next two data sets obtained from development testing of the Black Hawk helicopter and were made available by the Department of the Army. One set, denoted by RHYDR, is derived from a hydraulic subsystem and consists of 54 failures associated with 40 different failure modes. The second set, denoted by RT700, is derived from the engine and consists of 56 failures associated with 52 different failure modes.

B. Simulation Data

Inasmuch as actual reliability growth data is difficult to obtain, it is necessary to observe the operating characteristics of the different models on simulated data. Ideally, the simulation should be designed to

capture the salient features of real data but without conforming explicitly to any of the stochastic mechanisms underlying the models to be tested. The simulation employed in the present study is described below and was based roughly on the characteristics of the field data made available to the authors.

Five sets of independent simulated processes were constructed. In a given set, each simulation consists of the superposition of a number of independent homogeneous Poisson processes whose intensities are obtained by sampling from a gamma distribution with specified parameters. Each Poisson process which represents a different failure mode, is truncated after a certain number of failures have appeared. That number is determined by sampling independently from a Poisson distribution with a fixed mean.

The parameters determining each set of simulations are presented in Table 1. For example, data set 0 was generated by the superposition of 25 homogeneous Poisson processes whose intensities were sampled from a gamma distribution with a mean of 1.5 failures per 1000 hours and a mode of .786 failures per 1000 hours. (Those processes associated with the larger intensities correspond to failure modes which tend to appear early on in the reliability growth process.)

With real systems, some failure modes will only occur after a certain amount of non-stop operating time. For instance, a part may fail only after the engine has been running long enough for it to heat up to a certain

temperature. With this in mind, the simulations in data sets 0, 1, and 2 have been augmented by ten "lagged" failure modes, again modeled as independent homogeneous Poisson processes. However, these processes only begin after a threshhold of 80 consecutive hours without a failure has been reached. The intensities for these modes are obtained by sampling from a Gaussian distribution with fixed parameters. These are also displayed in Table 1. For data set 0, the mean is 5 failures per 1000 hours and the standard deviation 1 failure per 1000 hours. All lagged failure modes are truncated after three observed failures.

Data sets 3 and 4, although not augmented by lagged failures, involve a relatively large number of potential failure modes with low recurrence rates. All simulations were tracked until the first failure after 4000 hours.

4. Measures of Performance

The primary role of reliability growth models is to provide engineers with satisfactory predictions of future levels of reliability. We focus here on the accuracy of the model's short and long-term forecasts of instantaneous mean time between failures. Since it is difficult to determine absolute standards of performance, model assessments are carried out on a comparative basis.

One measure of the ability to predict interfailure times is given by the R2 statistic (Braun and Paine [1977]). It is simple modification of a statistic employed by Schafer, et al. [1975]. This statistic compares the average of three consecutive interfailure times to the estimated mean of the middle one of the three times. Even when reliability growth occurs, the observed X_i's can fluctuate greatly. The use of the average of three interfailure times hopefully diminishes some of the effect of the irregularities in the data. When using the R2 statistic, we only consider the failures which occur after a certain interval of time, when reliability growth should be more evident. It is usually the later interfailure times that we want to predict. Let

m = number of disjoint groups of three interfailure times after some starting time,

$\tilde{X}_j$ = average of the observed X_i's for the jth group, $j=1,\ldots,m$,

$E(\tilde{X}_j)$ = estimated mean interfailure time for the middle time in the jth group, $j=1,\ldots,m$, and

$$\bar{\tilde{X}} = \sum_{j=1}^{m} \tilde{X}_j/m.$$

The R2 statistic is given by

$$R2 = \frac{\sum_{j=1}^{m} (\tilde{X}_j - \tilde{E}(X_j))^2/(m-2)}{\sum_{j=1}^{m} (\tilde{X}_j - \tilde{\bar{X}})^2/(m-1)}.$$

A small value of R2 implies a good fit. Since R2 tests predictive value, it is proper, when estimating $E(X_j)$ for each j, to use the estimate derived from the data through time t_{j-1}.

A more severe test of the predictive power of a model is to estimate the parameters of the model using the data up to a certain test time and then to predict the level of the interfailure times at a later time, say 1000 hours later. Comparison of this prediction with the actual interfailure time level should give a good indication of the utility of the model. In this study, long-term forecasts of (current) reliability at 2500 and 3500 hours are obtained on the basis of observations up to 1500 hours. They are then compared with the observed instantaneous mean time between failures at 2500 and 3500 hours. These are actually estimated by taking the average of the 5 interfailure times straddling those time points.

Although of secondary interest, it is also possible to judge whether the distributional assumptions of the models are approximately satisfied by the date. If $F_i(\cdot)$ is the hypothesized cdf of the interfailure time, X_i, and $\hat{F}_i(\cdot)$ its estimate, then a plot of $\hat{F}_i(x_i)$ versus i should be roughly symmetric about the line $\hat{F}_i(x_i) = 0.5$ with no visible patterns. In addition, since the variates $u_i = \hat{F}_i(x_i)$ are approximately independent uniform on [0,1], 50% of the values should fall between .25 and .75. Clearly, many such tests may be devised.

5. Prediction with Failure Mode Models

As indicated in Section 2 above, prediction of future interfailure times with the Crow or urn models is relatively straightforward. However, since the intensity function $h_f(t)$ for the failure mode model depends on $N_1(t)$ and $N_2(t)$, prediction at some future time point is not simply accomplished. In this paper, we have employed the following procedure.

Suppose data is available until time t_o and prediction at time t_1 is required. Plot $\hat{h}_f(t)$ versus t at the jump points of the function. Ordinarily, $\hat{h}_f(t)$ will vary quite smoothly with t and, assuming that reliability growth eventually occurs, will be smoothly decreasing in t after a certain point. One could try to extrapolate $\hat{h}_f(\cdot)$ to t_1 by using a French curve or some such device. We have chosen instead to linearize the decreasing portion of the curve, by choosing appropriate reexpressions of t and $\hat{h}_f(t)$ (see Tukey [1977], Chapter 6), and extrapolate linearly on the transformed scale. The final prediction is obtained by transforming back to the original scales. Although this procedure is somewhat tedious in the context of a simulation study, it is quite practicable when a single data set is at hand.

6. Analysis of Simulated Data

A. R2 Statistics

The Crow, urn, and three versions of the FM model were applied to the ten simulations of set 0. The essential features of the batch of ten R2 statistics generated by each model are captured by the 5-number summaries (see Tukey [1977], Chapter 2) contained in Table 2 and displayed graphically in Figure 2. The urn model is clearly dominated by the other methods and seems prone to disproportionately large R2 values. Given the difficulty involved in obtaining the MLE's for this model, it was decided to eliminate it from future comparisons. On the whole, the FM models seem to perform slightly better than Crow's.

Simulations in data set 1 are characterized in general by fewer observed failures than those in data set 0. With less "data" available, the Crow model, which requires fewer parameters to be estimated, seems to perform slightly better overall. (See Table 3 and Figure 3.) This is particularly true with respect to the criterion of minimizing the chance of very large R2 values.

We attempted to improve the performance of all methods by smoothing the raw interfailure times before carrying out the parameter estimation. Smoothing was done either by moving averages of three or running medians of three. Unfortunately, the preprocessing of the data tended to degrade the performance of all the models. For comparison, Table 3 contains the 5-number summary for

R2's generated by the square-root version of the FM model applied to the interfailure times smoothed by running means.

For data set 2, with larger number of failures and an increased likelihood of observing lagged failures, the FM models seem to improve. (See Table 4 and Figure 4.) The square-root version is particularly noteworthy: it's median R2 value is nearly 40% less than that of Crow's and in a Q-Q plot (with Crow's quantiles along the abscissa) only one point, corresponding to the pair of largest order statistics, would lie above the 45^{o} line.

Another view of the comparison may be obtained from Figure 5 where the difference between the R2 statistics for the FM model and the Crow model for each simulation is plotted against the Crow R2 statistic. Here we see that in 17 out of 20 cases, the FM model proved superior to Crow's. The use of mean smoothing only succeeded in slightly reducing the largest order statistic.

The results for data sets 3 and 4 are analogous to those for data set 2, although the former have no lagged failures. The average observed number of failures is still comparable failure modes. The results are presented in Table 5 and 6, as well as Figures 6 through 9.

B. Long-term Prediction

We now focus on the quality of the models' long-term predictions as described in Section 4. Estimates of IMTBF at 2500 and 3500 hours were made for 10 simulations in each of data sets 3 and 4, on the basis of the information available at 1500 hours. Only the Crow model and the square-root version of the FM model were compared. The estimates and target

values are presented in Tables 7(a) and 8(a). One notes immediately the tremendous variability in the latter, despite the fact that they are actually averages of five consecutive interfailure times. On the whole, the FM-model predictions tend to be larger than the Crow predictions and display greater simulation-to-simulation variability.

The quality of the predictions can be assessed by consideration of Tables 7(b), 7(c), 8(b), and 8(c). They contain 5-number summaries of the errors of prediction and the relative errors of prediction. In interpreting the latter tables, it should be recalled that negative relative errors, corresponding to overprediction, are theoretically unbounded in magnitude while positive relative errors, corresponding to underprediction are bounded above by unity. Thus, the FM-model which occasionally overpredicts small observed values earns large negative relative errors. On the other hand, the Crow model tends to considerably underpredict future interfailure times: at 3500 hours, for example, Crow's median relative error exceeds + 50% for both data sets. Predictions of the FM-model are more nearly unbiased and somewhat superior to those of Crow's model. Although both sets of predictions could use considerable improvement, it should be kept in mind that at 3500 hours one is extrapolating over a time interval longer than the period of observation. Such extrapolation is necessarily prone to large errors.

C. Goodness-of-Fit

Figures 10 and 11 display plots of $\hat{F}_i(X_i)$ versus i for the FM and Crow models from a simulation in data set 3. Both models seem to fit the data well and this appears to be true in general. These results underline the danger in inferring from goodness-of-fit tests the adequacy of long-term predictions.

7. Analysis of Field Data

Smoothed plots of CMTBF versus Test time for the RT 700 and RHYDR data sets are displayed in Figures 12 and 13. Only the RT 700 system seems to display sustained reliability growth. In Table 9 we present the R2 statistics of the Crow model and the FM models for all four data sets. Except for the RHYDR set, where it performs abysmally, the Crow model is superior. Plots of $\hat{F}_i(X_i)$ versus i for the different models do not differ appreciably. Figures 14 and 15 provide one example.

Long term prediction was carried out for the GE data only. The Crow model was compared to the square-root version of the FM model. The former was more accurate at 2500 hours, the latter at 3500 hours.

8. Conclusions

The basic conclusion of this study seems to be that the family of FM models is deserving of further consideration. Although these models require one more parameter than does Crow's, significant improvements in forecasting accuracy result. One drawback to the FM family is the effort required to produce the long-range forecasts. However, this should not be a bar to practical implementation. Validation in field testing will determine whether the extra effort is merited.

We have also determined that preprocessing of the raw interfailure times does not improve parameter estimates or lead to better forecasts.

Future research should be directed at determining stochastic properties of the FM models, developing estimates of standard errors of prediction and further analysis of field data.

Acknowledgments

The authors would like to thank Mr. R. C. Kroeger of the General Electric Company for making available the GE data set, and Mr. R. J. Neff of the U.S. Army Aviation Research and Development Command for making available the RT700 and RHYDR data sets. Figures 12 and 13 are derived from supplementary material provided by Mr. Neff.

References

Braun, H. I. and Paine, J. M., (1977). " A Comparative Study of Models for Reliability Growth," Department of Statistics, Princeton University, Technical Report No. 126, Series 2.

Crow, L. H., (1975). "Reliability Analysis for Complex Repairable Systems," U.S. Army Material Systems Analysis Activity, Technical Report No. 138, Aberdeen Proving Ground, Maryland.

Duane, J. T., (1964). "Learning Curve Approach to Reliability Monitoring," IEEE Transactions on Aerospace, 2. pp. 563-566.

Schafer, R. E., Sallee, R. B. and Torrez, J. D., (1975). "Reliability Growth Study," Hughes Aircraft Company, Rome Air Development Center, Technical Report-75-253, Griffis Air Force Base, New York.

Tukey, J. W., (1977). Exploratory Data Analysis, Addison-Wesley Publishing Company: Reading, MA.

Appendix

A. Crow Model

The intensity function for the nonhomogeneous Poisson process underlying the Crow model is $h_c(t) = kbt^{b-1}$, $t>0$.

Suppose that n failures are observed, the last one occurring at time t(n). Then the maximum likelihood estimates of k and b are found to be (Crow [1975]):

$$\hat{k} = n/[t(n)]^{\hat{b}}$$

and

$$\hat{b} = n/\sum_{i=1}^{n-1} \log[t(n)/X_i]$$

B. Failure Mode Models

The intensity function for the failure mode model satisfies the relation $\ln(h_f(t)) = d_o + d_1 \ln(N_1(t) + 1) + d_2\, g(N_2(t) + 1)$. For the case that $g(\cdot)$ is the square root, we have

$$h_f(t) = \exp\{d_o + d_2 \sqrt{N_2(t) + 1}\ \}\ (N_1(t) + 1)^{d_1}$$

Let $M_{1i} = N_1(t_{i-1}) + 1$ and $M_{2i} = N_2(t_{i-1}) + 1$.

Then,

$$\ln L(X_1, X_2, \ldots, X_n)$$

$$= \sum_{i=1}^{n} - h_f(t_{i-1})\, X_i \ln h_f(t_{i-1})$$

$$= \sum_{i=1}^{n} \{d_o + d_2 \sqrt{C_{2i}} + d_i \ln(C_{1i}) - \exp(d_o + d_2 \sqrt{C_{2i}})\, C_{1i}^{d_1} X_i\}.$$

Setting the derivatives of ln L with respect to d_0, d_1, and d_2 equal 0, yields the equations:

$$d_o = \ln n - \ln \sum_{i=1}^{n} y_i \quad , \qquad \text{(B1)}$$

$$\sum_{i=1}^{n} \ln C_{1i} = [n/\sum_{i=1}^{n} y_i] \sum_{i=1}^{n} y_i \ln C_{1i}, \qquad \text{(B2)}$$

$$\sum_{i=1}^{n} \sqrt{C_{2i}} = [n/\sum_{i=1}^{n} y_i] \sum_{i=1}^{n} y_i \sqrt{C_{2i}} \ , \qquad \text{(B3)}$$

where

$$y_i = \exp \ (d_2 \sqrt{C_{2i}}) \ C_{1i}^{d_1} X_i.$$

Equations (B2) and (B3) can be solved by ordinary Newton - Raphson iteration and the resulting solutions for d_1 and d_2 employed in (B1) to obtain an estimate for d_o.

C. Urn Model

Suppose that the sequence of observed interfailure times is $\{X_i = j_i \delta\}$ $i = 1, 2, \ldots, n$, where j_i is some integer. Let $k = C/N$ and $a_i = i/N$ $(i = 1, 2, \ldots, n)$. Then

$$\Pr \{X_{i+1} = j_{i+1} \delta\} = [(a_i + k)/(1 + k)]^{j_{i+1}-1} (1 - a_i)/(1 + k).$$

Hence,

$$\log L_{N,K}(X_1, X_2, \ldots, X_n)$$

$$= \sum_{i=o}^{n-1} \{(j_{i+1} - 1) \ \log [(a_i+k)/(1+k)] + \log [(1-a_i)/(1+k)]\}.$$

Setting $\partial \log L/\partial k$ equal 0, yields the equation

$$\sum_{i=o}^{n-1} [(j_{i+1} - 1) \ (1-a_i)/(a_i + k)] = n. \qquad \text{(C1)}$$

Setting $\partial \log L/\partial k$ equal 0, yields the equation

$$\sum_{i=0}^{n-1} \frac{i}{N+i} = \sum_{i=0}^{n-1} i(j_{i+1}-1)/(i + kN). \qquad (C2)$$

The simplest approach to finding solutions to (C1) and (C2) seems to be to search over a grid of values of k and N. However, this consumes a fair amount of computer time and the process does not always converge to acceptable values of k and N.

			PRIMARY FAILURE MODES				LAGGED MODES	
DATA SET	# SIMULATIONS	AVERAGE # FAILS	# MODES	INTENSITY MEAN	INTENSITY MODE	TRUNCATION MEAN	INTENSITY MEAN	INTENSITY S.D.
0	10	78.3	25	1.5	.786	3	5	1
1	20	58.1	25	1	.80	2	2	.5
2	20	63.4	25	2	.795	2	2	.5
3	20	70.0	40	1.5	.786	2	-	-
4	20	72.6	80	1.5	.786	1	-	-

Table 1. Simulation Parameters.

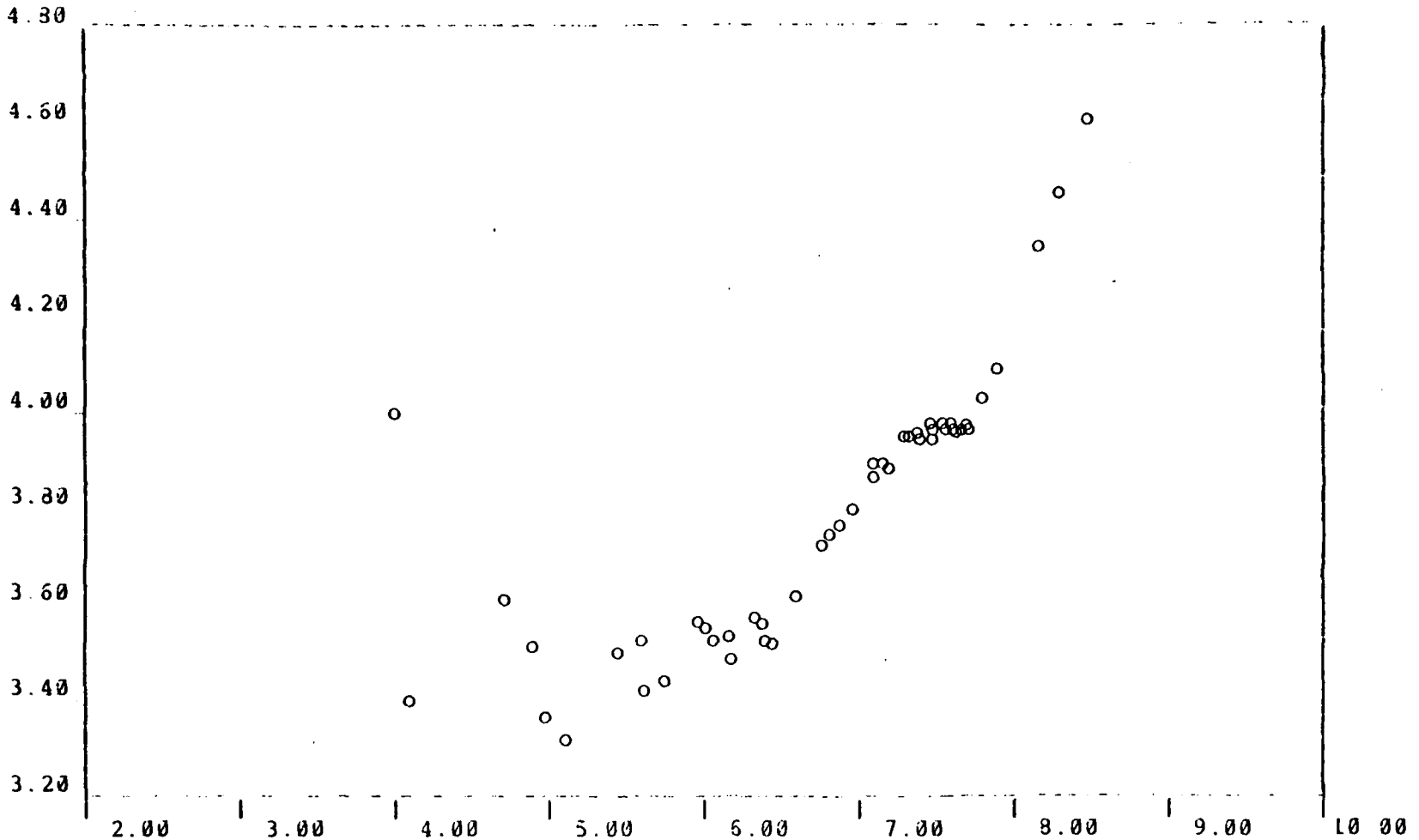


Figure 1a. Simulation data with no lagged failure modes.
Plot of ln (CMTBF) vs. ln (t).

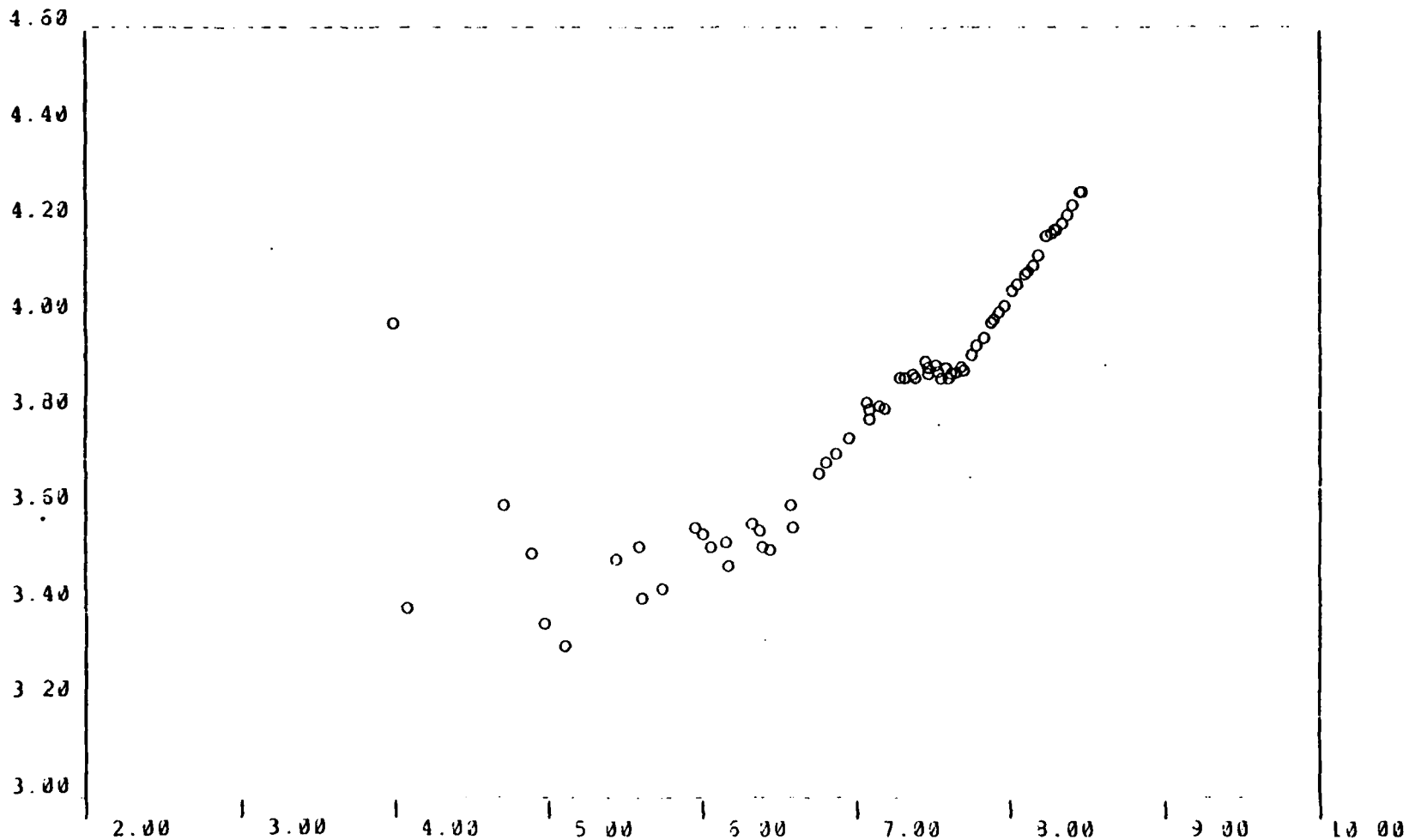


Figure 1b. Simulation data with lagged failure modes.
Plot of ln (CMTBF) vs. ln (t).

MODEL	n	0%	25%	50%	75%	100%
1	10	0.9400	1.080	1.330	1.510	2.890
2	10	1.150	1.310	1.435	2.370	4.770
3	10	0.8400	0.9900	1.165	1.200	2.990
4	10	0.6300	0.7700	1.010	1.640	2.530
5	10	0.6500	0.9100	1.255	1.940	2.590

(1): CROW

(2): URN

(3): FM, $\ln(N_2+1)$

(4): FM, $\sqrt{(N_2+1)}$

(5): FM, N_2+1

Table 2. Data Set 0. 5-number summary of R2 statistics.

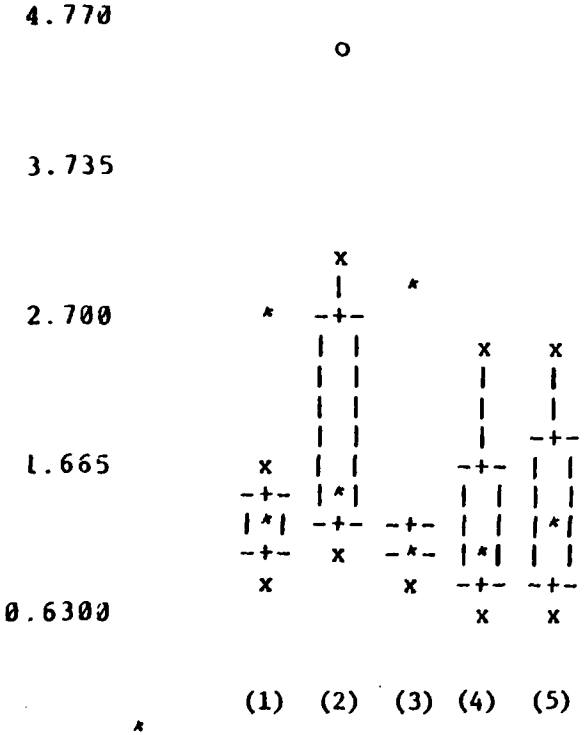


Figure 2. Data Set 0. Comparison plot based on Table 2.

MODEL	n	0%	25%	50%	75%	100%
1	20	0.8600	1.230	1.330	1.485	1.710
2	20	0.7200	1.000	1.190	1.335	1.710
3	20	0.6000	1.090	1.255	1.455	2.410
4	20	0.5700	1.140	1.315	1.680	3.470
5	20	0.6200	0.960	1.285	1.455	2.760

(1): CROW

(2): FM, $\ln(N_2+1)$

(3): FM, $\sqrt{(N_2+1)}$

(4): FM, N_2+1

(5): FM, $\sqrt{N_2+1}$; Data smoothed by moving average.

Table 3. Data Set 1. 5-number summary of R2 statistics.

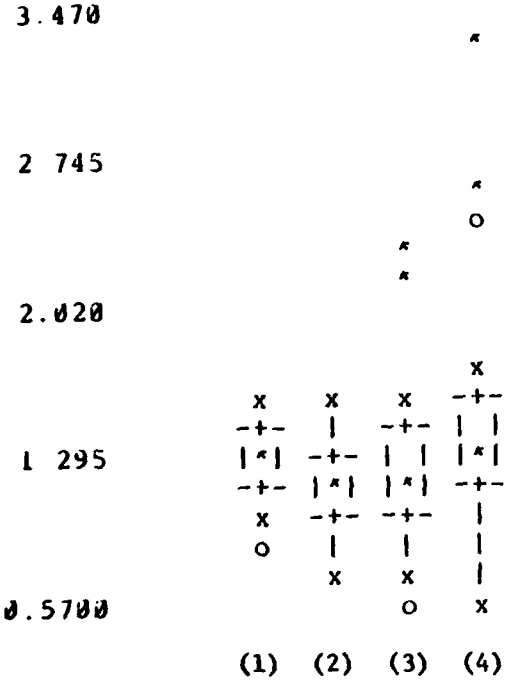


Figure 3. Data Set 1.

Comparison plot based on Table 3.

MODEL	n	0%	25%	50%	75%	100%
1	20	0.8600	1.020	1.240	1.510	1.710
2	20	0 5700	0.8400	1.060	1.300	1.630
3	20	0.5300	0.6150	0.8600	1.190	2.080
4	20	0.4600	0.6700	0.9700	1.615	3.050
5	20	0.5400	0.6150	0.8650	1.200	1.890

See Notes for Table 3.

Table 4. Data Set 2. 5-number summary of R2 statistics.

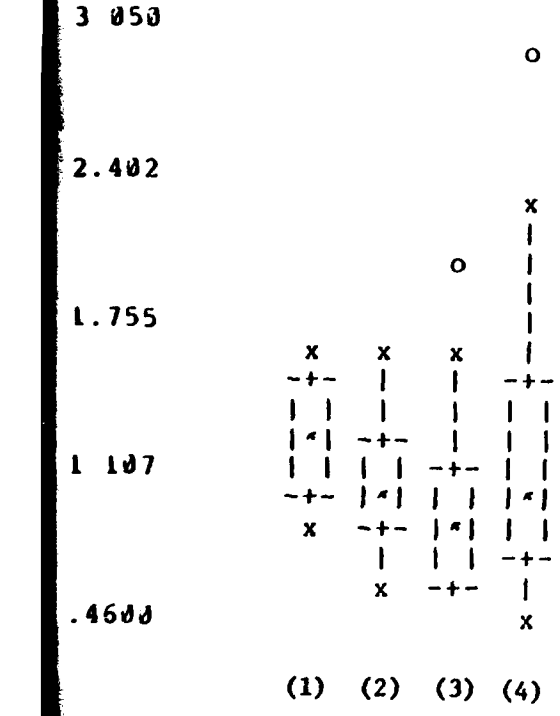


Figure 4. Data Set 2.

Comparison plot based on Table 4.

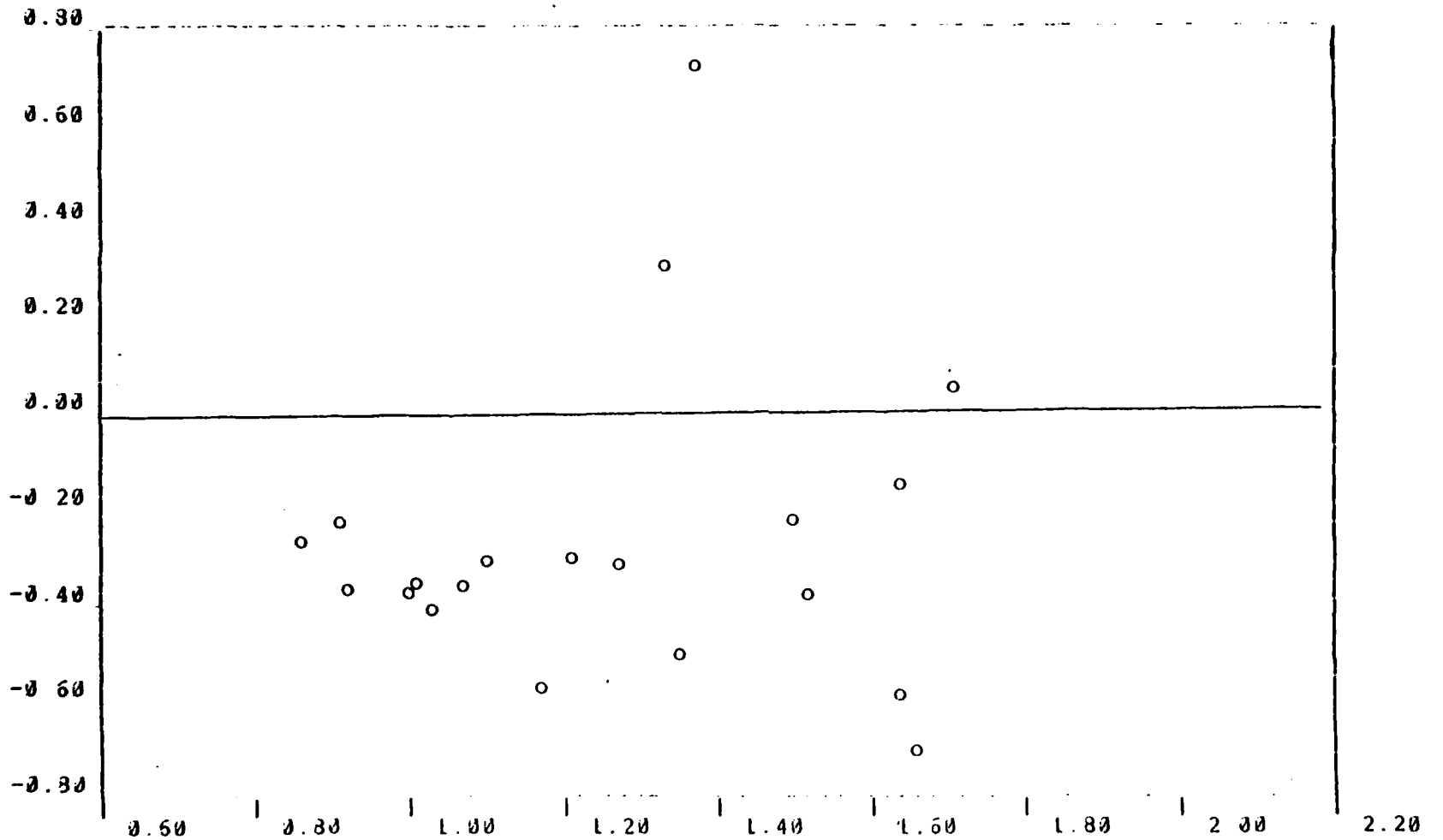


Figure 5. Data Set 2. R2 statistics. FM-CROW vs. CROW.

MODEL	n	0%	25%	50%	75%	100%
1	20	0.8500	1 020	1.080	1 125	1.280
2	20	0 5600	0 8400	0 9100	1.075	1.210
3	20	0.4200	0 6300	0 8450	1.010	1.220
4	20	0 3000	0.5800	0 8950	1.025	1.410
5	20	0.3700	0.6350	0.8550	1.005	1.190

See Notes for Table 3.

Table 5. Data Set 3. 5-number summary of R2 statistics.

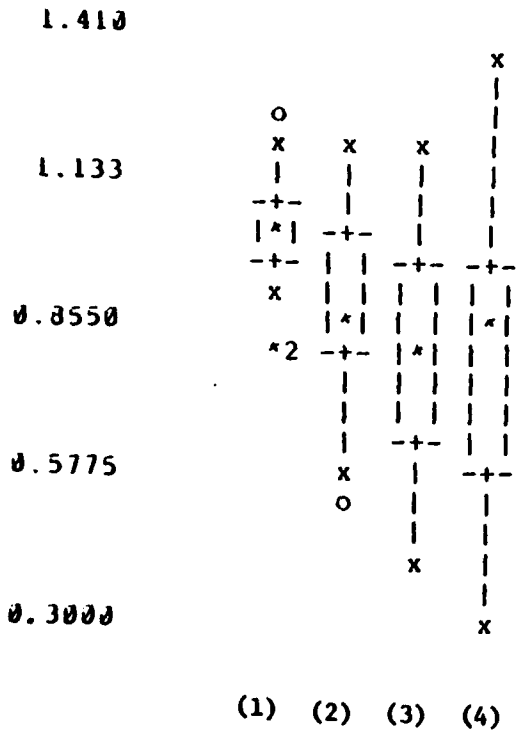


Figure 6. Data Set 3.
Comparison plot based on Table 5

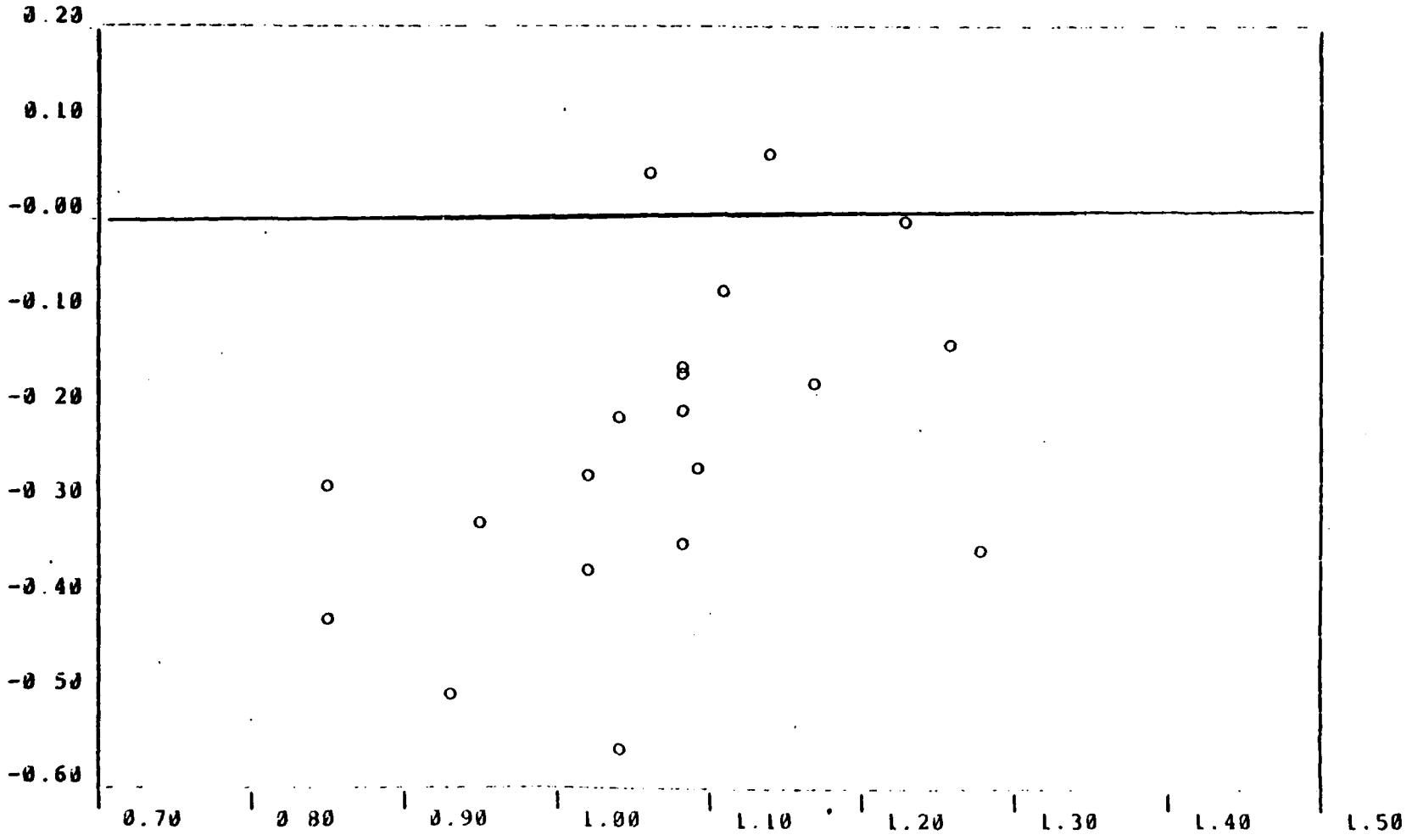


Figure 7. Data Set 3. R2 statistics. FM-CROW vs. CROW.

MODEL	n	0%	25%	50%	75%	100%
1	20	0.8100	0.9200	0.9850	1.030	1.170
2	20	0.4500	0.5450	0.6300	0.9800	1.120

(1): CROW

(2): FM, $\sqrt{(N_2+1)}$

Table 6. Data Set 4. 5-number summary of R2 statistics.

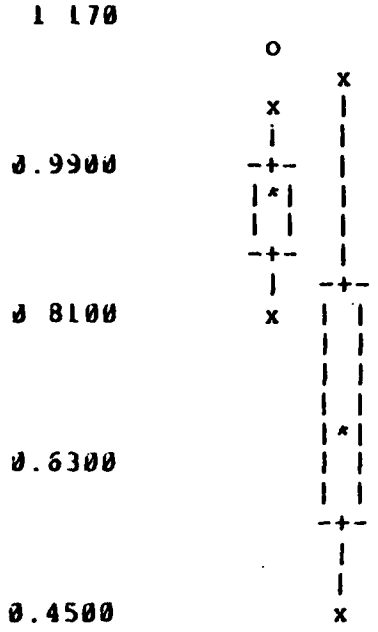


Figure 8. Data Set 4.

Comparison plot based on Table 6.

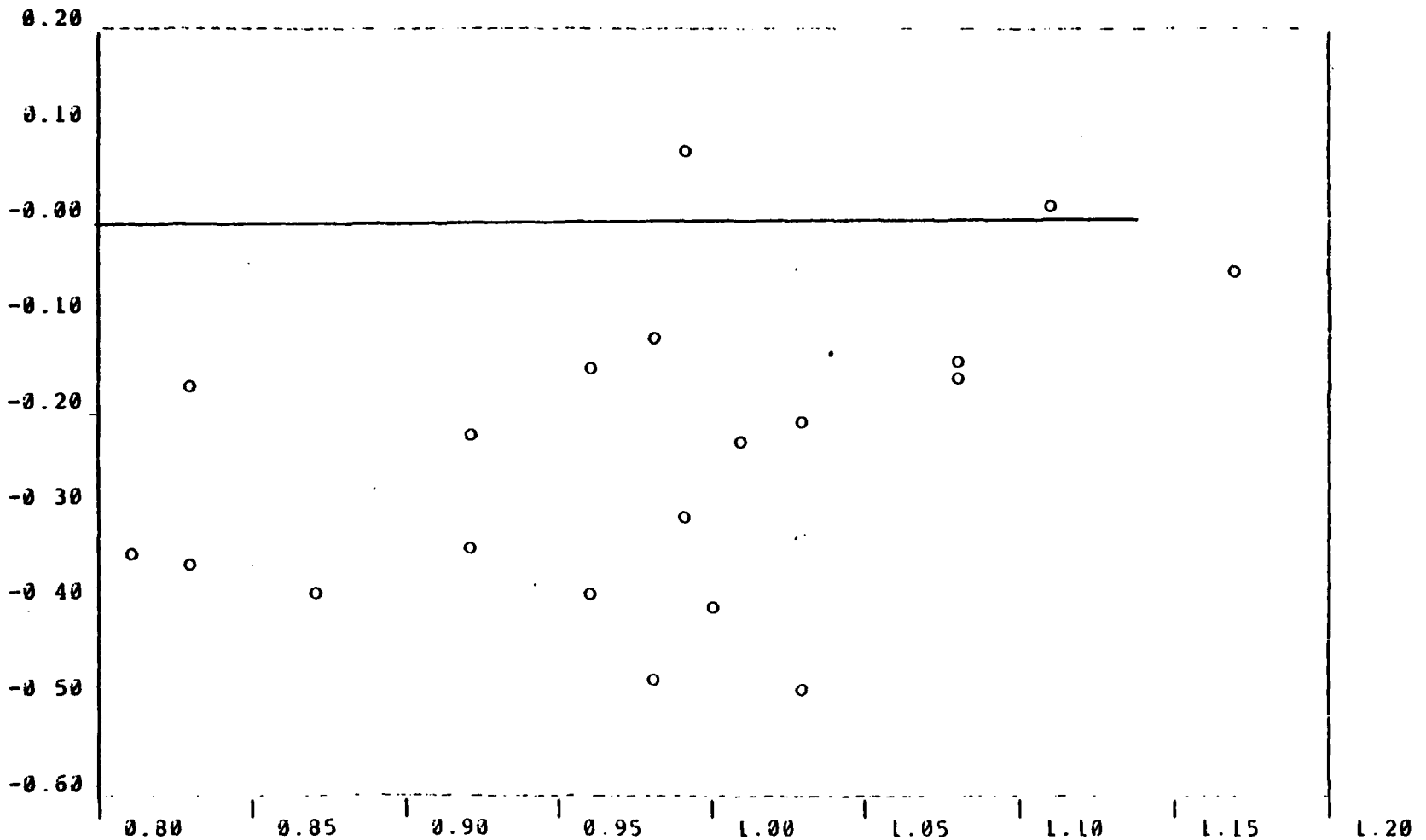


Figure 9. Data Set 4. R2 statistics. FM-CROW vs. CROW.

2500 HOURS			3500 HOURS		
PREDICTIONS		OBSERVED	PREDICTIONS		OBSERVED
FM, $\sqrt{N_2+1}$	CROW		FM, $\sqrt{N_2+1}$	CROW	
60.5	33.9	99.8	95.0	54.6	171.1
52.8	36.5	26.4	90.3	64.6	233.9
49.0	40.0	67.9	88.5	54.4	137.6
253.6	70.9	55.6	166.9	81.6	262.5
90.4	49.2	102.2	239.4	95.0	312.4
108.2	39.7	120.4	154.2	75.2	64.9
73.7	55.7	71.3	143.7	88.2	123.9
67.6	39.4	44.3	94.2	56.2	94.2
64.3	37.5	62.7	93.3	67.0	58.9
89.4	43.3	32.4	69.5	48.3	70.0

(a) Raw Data

MODEL	TIME	0%	25%	50%	75%	100%
FM, $\sqrt{N+1}$	2500	-198.0	-26.4	-2.0	12.2	39.3
CROW	2500	- 15.3	-10.1	15.6	53.0	80.7
FM, $\sqrt{N+1}$	3500	- 89.3	-19.8	24.8	76.1	143.6
CROW	3500	- 10.3	21.7	60.6	169.3	217.4

(b) Errors of Prediction. 5-number summaries.

FM, $\sqrt{N+1}$	2500	-3.56	-1.76	-.03	.12	.39
CROW	2500	- .38	- .28	.31	.52	.67
FM, $\sqrt{N+1}$	3500	-1.38	- .16	.12	.36	.61
CROW	3500	- .16	.29	.50	.69	.72

(c) Relative Errors of Prediction. 5-number summaries.

Table 7. Data Set 3. Long-term Forecasts of IMTBF.

2500 HOURS			3500 HOURS		
PREDICTIONS		OBSERVED	PREDICTIONS		OBSERVED
FM, $\sqrt{N_2+1}$	CROW		FM, $\sqrt{N_2+1}$	CROW	
56.4	33.5	53.7	70.4	50.7	180.0
93.9	42.8	46.6	124.1	58.0	100.6
132.0	77.2	48.2	70.2	90.2	121.9
112.0	47.6	148.3	174.8	76.7	151.7
123.5	41.8	83.9	163.2	80.2	275.8
97.7	51.1	112.4	181.4	82.4	268.8
74.3	56.8	83.7	191.2	102.0	365.8
91.6	58.1	77.8	135.5	82.9	223.9
77.9	44.8	32.1	93.3	57.7	147.7
123.6	63.8	154.5	206.7	97.6	206.5

(a) *Raw Data*

MODEL	TIME	0%	25%	50%	75%	100%
FM, $\sqrt{N+1}$	2500	-83.8	-45.8	-8.2	14.7	36.3
CROW	2500	-29.0	3.8	23.5	61.3	100.7
FM, $\sqrt{N+1}$	3500	-23.5	.2	70.9	109.6	174.6
CROW	3500	31.7	75.0	119.2	186.4	263.8

(b) Errors of Prediction. 5-number summaries.

FM, $\sqrt{N+1}$	2500	-1.74	-1.02	-.11	.13	.24
CROW	2500	- .60	.08	.34	.55	.68
FM, $\sqrt{N+1}$	3500	- .23	0	.38	.42	.61
CROW	3500	.26	.49	.62	.71	.72

(c) Relative Errors of Prediction. 5-number summaries.

Table 8. Data Set 4. Long-term Forecasts of IMTBF.

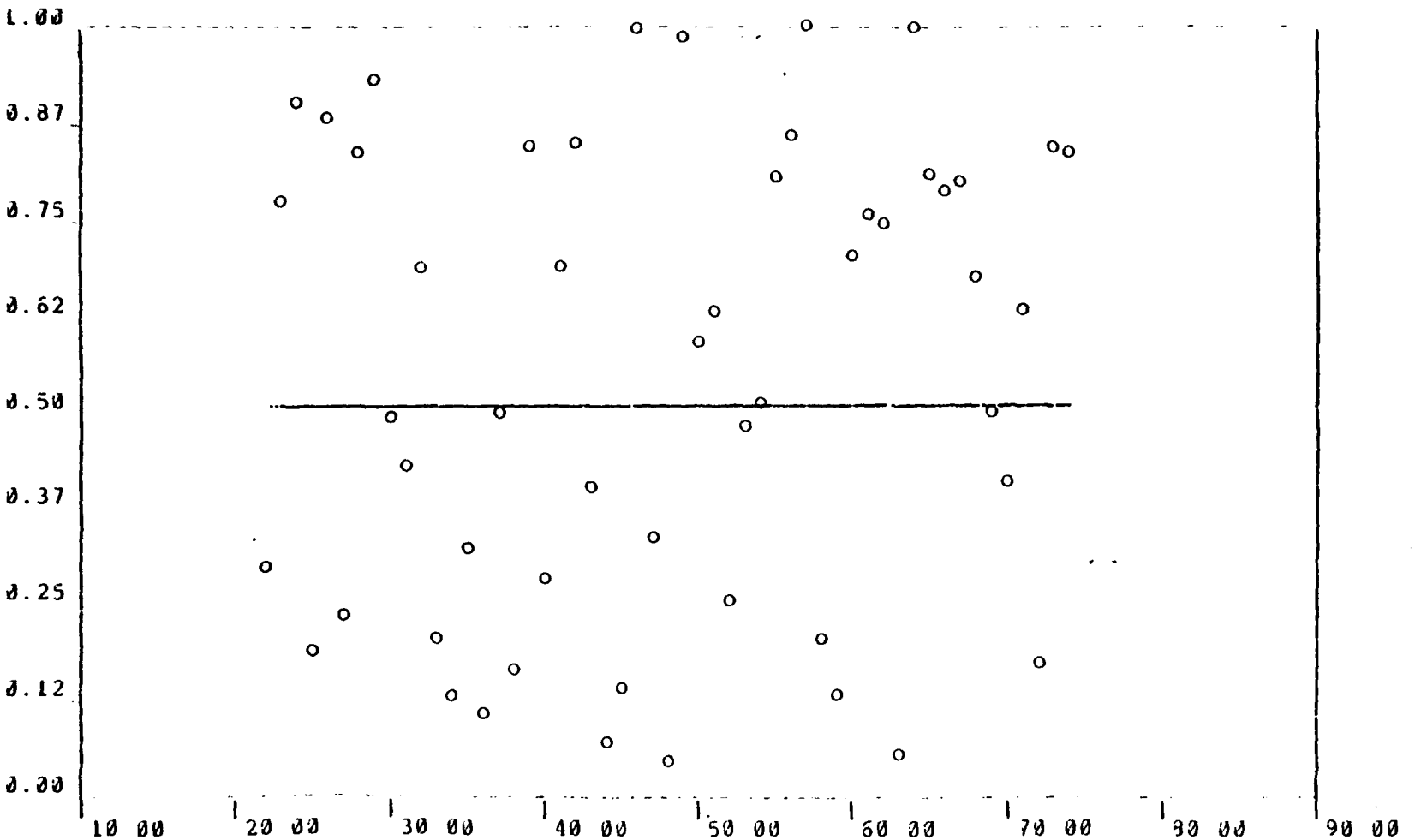


Figure 10. Data Set 3 Simulation. $\hat{F}_i(X_i)$ vs. i for FM, $\sqrt{N+1}$ Model.

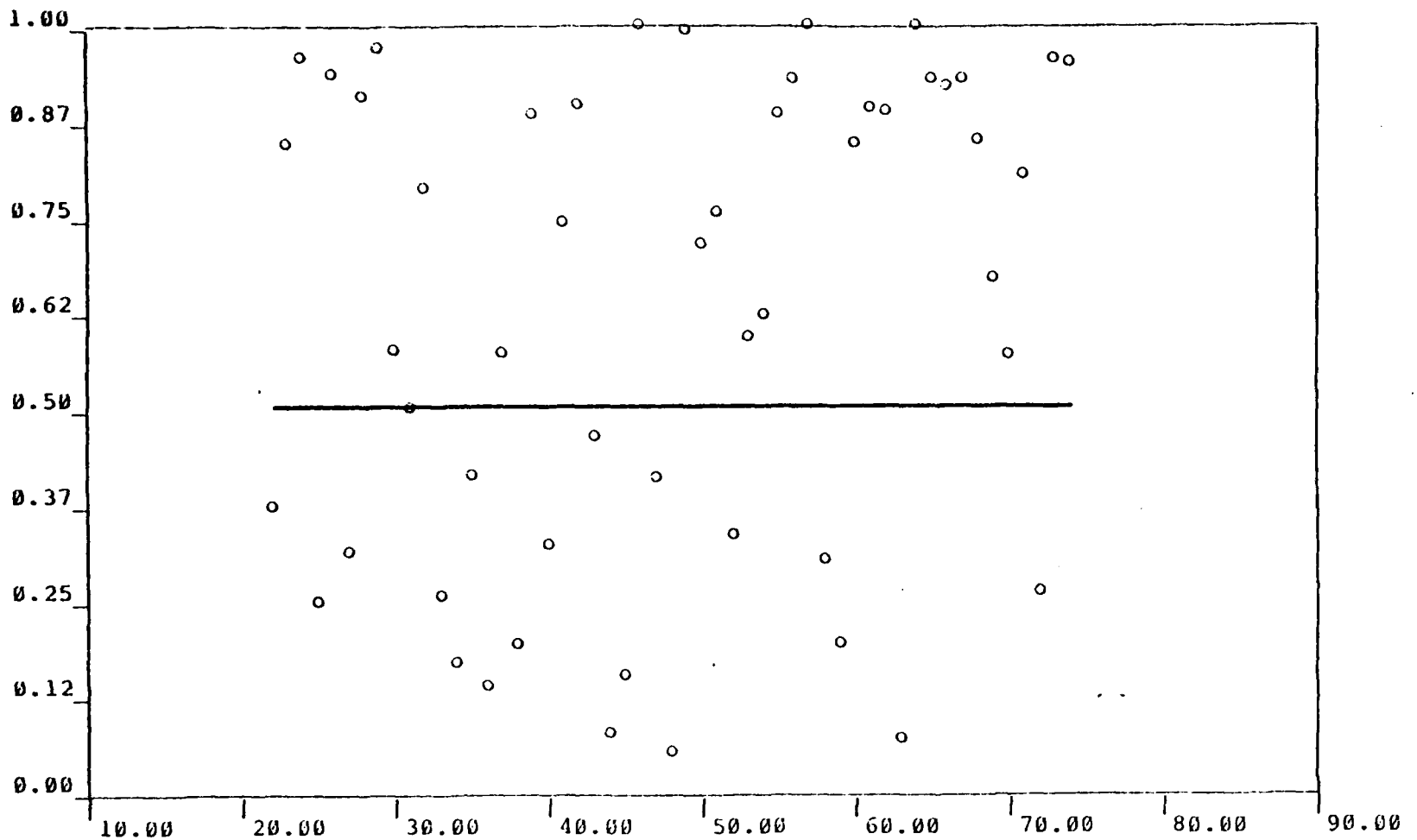


Figure 11. Data Set 3 Simulation. $\hat{F}_i(X_i)$ vs. i for Crow Model.

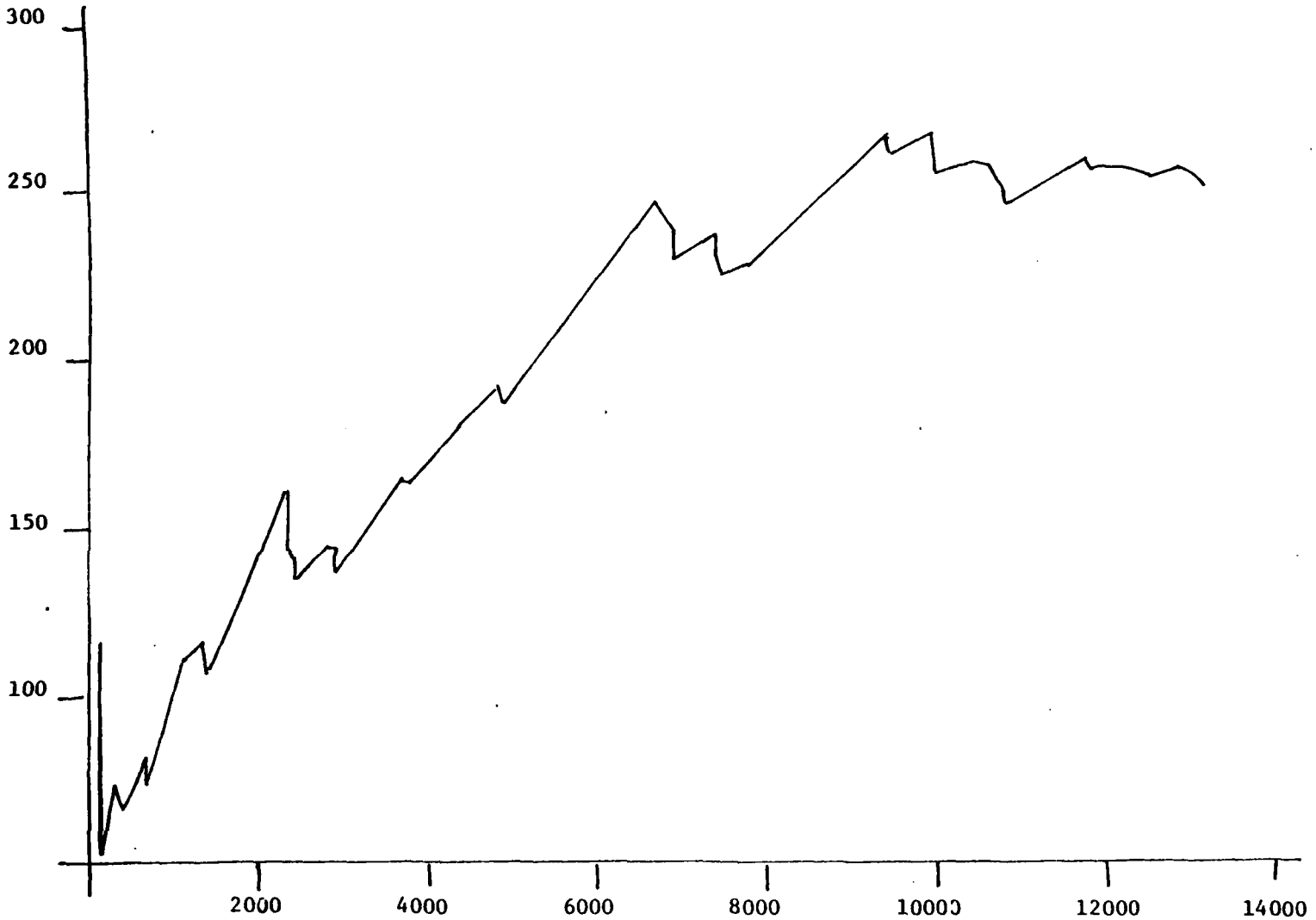


Figure 12. RT700 Data Plot of CMTBF vs. Test time.

Figure 13. RHYDR Data. Plot of CMTBF vs. Test time.

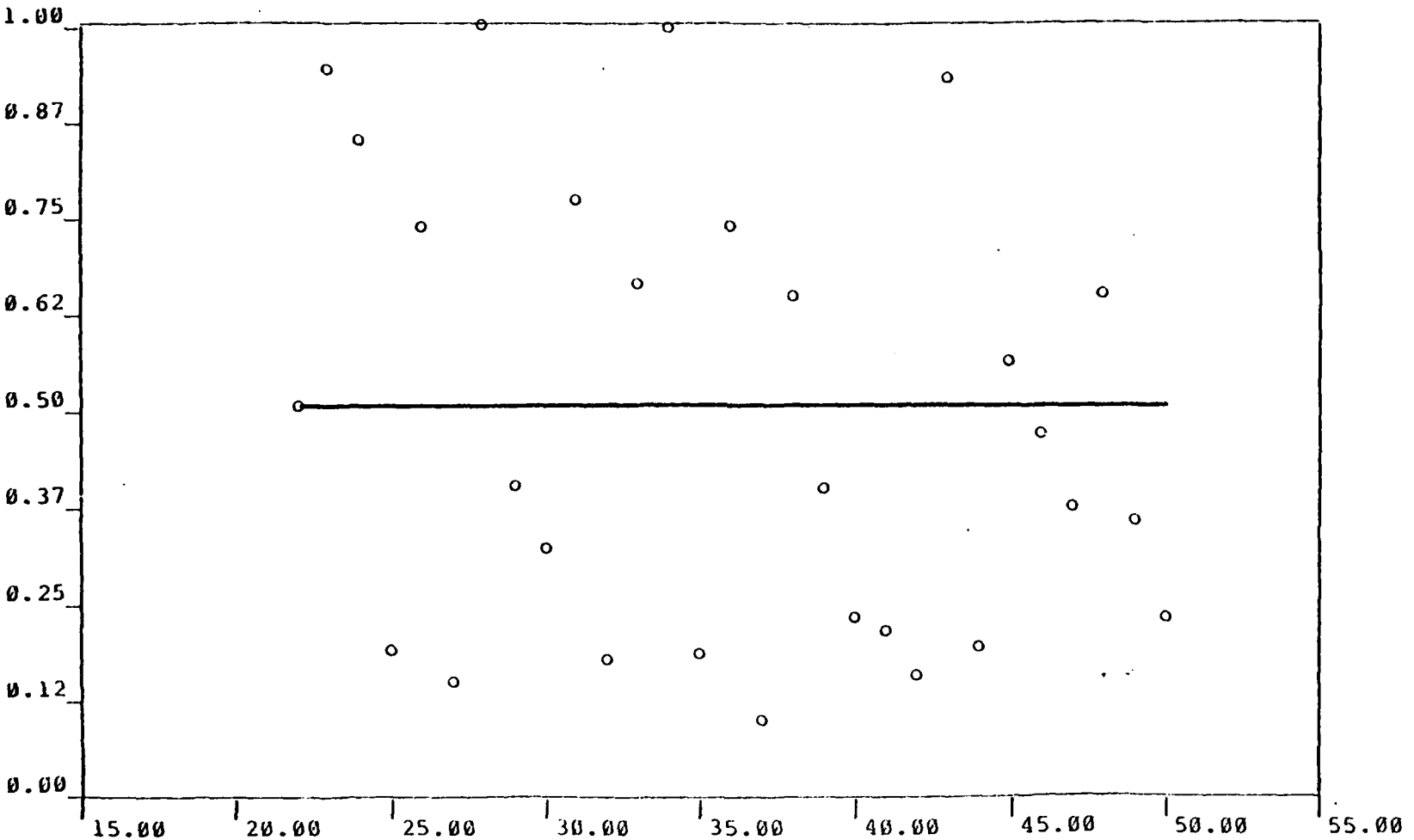


Figure 14. RT700 Data. Plot of $\hat{F}_i(X_i)$ vs. i for Crow Model.

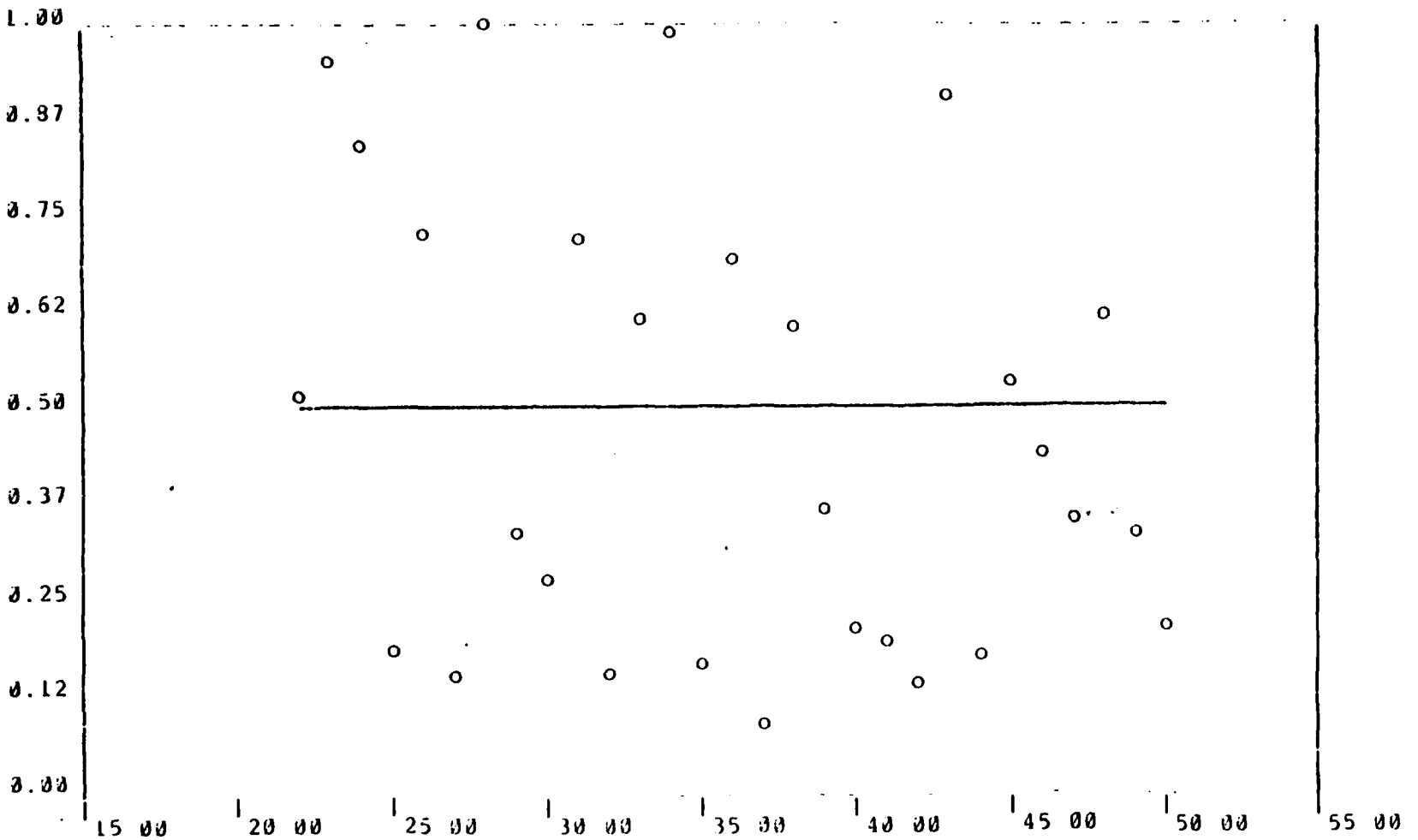


Figure 15. RT700 Data. Plot of $\hat{F}_i(X_i)$ vs. i for FM, $\sqrt{N+1}$.